

2020-10-27

# Next-generation sequencing of 35 RHD variants in 16253 serologically D pregnant women in the Finnish population

Tammi, SM

<http://hdl.handle.net/10026.1/16637>

---

10.1182/bloodadvances.2020001569

Blood Advances

American Society of Hematology

---

*All content in PEARL is protected by copyright law. Author manuscripts are made available in accordance with publisher policies. Please cite only the published version using the details provided on the item record or document. In the absence of an open licence (e.g. Creative Commons), permissions for further reuse of content should be sought from the publisher or author.*

# Next-generation sequencing of 35 *RHD* variants in 16 253 serologically D<sup>−</sup> pregnant women in the Finnish population

Silja M. Tammi,<sup>1,\*</sup> Wajnat A. Tounsi,<sup>2,3,\*</sup> Susanna Sainio,<sup>1</sup> Michele Kiernan,<sup>2</sup> Neil D. Avent,<sup>2</sup> Tracey E. Madgett,<sup>2</sup> and Katri Haimila<sup>1</sup>

<sup>1</sup>Finnish Red Cross Blood Service, Helsinki, Finland; <sup>2</sup>School of Biomedical Sciences, Faculty of Health, University of Plymouth, Plymouth, United Kingdom; and <sup>3</sup>School of Medical Technology, King Abdulaziz University, Jeddah, Saudi Arabia

## Key Points

- Eight novel *RHD* variant alleles have been detected in the Finnish population following NGS of the whole *RHD* gene.
- NGS is a powerful method for genotyping the highly polymorphic *RHD* gene.

Fetal *RHD* screening for targeted routine antenatal anti-D prophylaxis has been implemented in many countries, including Finland, since the 2010s. Comprehensive knowledge of the *RHD* polymorphism in the population is essential for the performance and safety of the anti-D prophylaxis program. During the first 3 years of the national screening program in Finland, over 16 000 samples from RhD<sup>−</sup> women were screened for fetal *RHD*; among them, 79 samples (0.5%) containing a maternal variant allele were detected. Of the detected maternal variants, 35 cases remained inconclusive using the traditional genotyping methods and required further analysis by next-generation sequencing (NGS) of the whole *RHD* gene to uncover the variant allele. In addition to the 13 *RHD* variants that have been previously reported in different populations, 8 novel variants were also detected, indicating that there is more variation of *RHD* in the RhD<sup>−</sup> Finnish population than has been previously known. Three of the novel alleles were identified in multiple samples; thus, they are likely specific to the original Finnish population. National screening has thus provided new information about the diversity of *RHD* variants in the Finnish population. The results show that NGS is a powerful method for genotyping the highly polymorphic *RHD* gene compared with traditional methods that rely on the detection of specific nucleotides by polymerase chain reaction amplification.

## Introduction

Immunization against the RhD antigen is the major cause of severe hemolytic disease of the fetus and newborn (HDFN).<sup>1</sup> Thus, a comprehensive anti-D prophylaxis program is required to prevent immunizations. In addition to the postnatal prophylaxis implemented in the 1960s, routine antenatal anti-D prophylaxis is also now offered by many countries during the third trimester of pregnancy.<sup>2</sup> The development of noninvasive fetal *RHD* genotyping has enabled the targeting of routine antenatal anti-D prophylaxis only to women carrying an RhD<sup>+</sup> child, which prevents up to 40% of pregnant RhD<sup>−</sup> women carrying an RhD<sup>−</sup> fetus from receiving unnecessary prophylaxis.<sup>3</sup>

Finland implemented national fetal *RHD* screening for targeted anti-D prophylaxis for all RhD<sup>−</sup> pregnant women in February 2014. Fetal *RHD* screening is offered to all women who have been typed serologically as D<sup>−</sup>. Pregnant women with partial D variants and certain weak D variants, as well as those who have inconsistent phenotyping and genotyping results, are also considered RhD<sup>−</sup>, so that they are included in the prophylaxis program, due to the risk of immunization.

A maternal *RHD* variant can be suspected in the fetal *RHD* screening based on a low cycle-threshold value in the real-time polymerase chain reaction (PCR) assay, indicative of the high amount of maternal

Submitted 29 January 2020; accepted 11 August 2020; published online 14 October 2020. DOI 10.1182/bloodadvances.2020001569.

\*S.M.T. and W.A.T. contributed equally to this study.

The novel *RHD* variants reported in this article have been deposited in the GenBank database (accession numbers MN365996, MN365995, MN366000, MN365998, MN366001, MN365997, MN365999, and MN366002).

© 2020 by The American Society of Hematology

DNA present. The PCR-based genotyping methods primarily used for the identification of maternal variants rely on the detection of specific nucleotides by PCR amplification, so that only a limited number of known variants can be identified. Novel variants outside of the targeted regions therefore remain undetermined.<sup>4,5</sup>

The development of massively parallel high-throughput next-generation sequencing (NGS) has offered new possibilities for blood group genotyping.<sup>4,6-8</sup> NGS can be efficiently applied for *RHD* genotyping, although the high level of homology between the *RHD* and *RHCE* genes is a challenge.<sup>4,9-13</sup> The NGS approach can sometimes offer the advantage of detecting the *cis/trans* linkage of polymorphisms, revealing heterozygous compound mutations, which cannot be identified in the conventional Sanger sequencing due to signal overlay in the electropherogram.

During the first 3 years of the national Finnish screening program, 16 253 samples of D<sup>-</sup> pregnant women were screened for fetal *RHD*. From this cohort, 79 maternal *RHD* variants were detected. Of the detected variants, 35 samples needed further scrutinizing because the sequence-specific primer PCR (SSP-PCR) method indicated the presence of a normal *RHD* gene although it did not correspond to the phenotype of the samples. In this study, these samples were genotyped by NGS for the complete *RHD* gene to identify the causative mutations for the negative or inconclusive RhD-phenotyping results.

## Materials and methods

All 16 253 samples were from pregnant RhD<sup>-</sup> women who participated in the national screening program for HDFN in Finland between February 2014 and December 2016. The sample collection and processing was performed as described by Haimila et al.<sup>14</sup>

The study was approved by the Research Ethics Committee of Helsinki University Hospital Women's clinic (271/13/03/03/2014) and by the National Institute for Health and Welfare (1365/5.05.01/2014).

### RhD phenotyping

All samples were phenotyped for RhD using a solid-phase microplate system by Olympus (PK 7300 automate; Beckman Coulter). Two monoclonal anti-D reagents were used, both in 2 dilutions: the clone LDM3/ESD1 (Alba Bioscience Limited) and the clone TH-28/MS-26 (Bioscot; Millipore). If the premier method did not give a clear positive (n.b.: RhD<sup>+</sup> samples were not included in this study) or negative result, an IH-1000 instrument (Bio-Rad) was also used to test a sample for RhD according to the manufacturer's instructions. The anti-D reagent used in the IH-1000 instrument was LHM 59/20 (LDM3) in ABO/D reverse grouping gel cards (Bio-Rad).

All samples giving a clear negative result, or a weak positive reaction ( $\leq 2+$ ) and a genotype not corresponding to the phenotype leaving the RhD result inconclusive, were designated as RhD<sup>-</sup>, meaning that women were included in the anti-D prophylaxis program to prevent possible anti-D immunization.

To estimate the amount of D antigen on red blood cells, an absorption-elution test was performed. For anti-D adsorption, 700  $\mu$ L of washed red blood cells (RBCs) were incubated for 1 hour at 37°C with an equal volume of anti-D (anti-D [Rh1], monoclonal, P3x61; Diagnost). The cells were washed 5 times

(Titriplex III 0.09 mol/L, disodium hydrogen phosphate dihydrate 0.026 mol/L, NaCl 0.148 mol/L, pH 6.8). Elution was effected by an acid-elution buffer, pH 2.8, and neutralized by an alkaline buffer, pH 10.4. The eluate and the last washed supernatant were used for the direct antiglobulin test against D<sup>+</sup> (CcDEe) reagent RBCs. Serial dilutions of eluate in phosphate-buffered saline were prepared. Next, 60  $\mu$ L of 4% D<sup>+</sup> RBCs were added to each tube, incubated 15 minutes at 37°C, and centrifuged at 500g for 10 seconds. The agglutination was read macroscopically and the reciprocal of the highest eluate dilution that gave a 1+ agglutination reaction was interpreted as the titer.

### RHD genotyping

All 16 253 samples were screened for fetal *RHD* as described by Haimila et al.<sup>14</sup> In 79 samples, a maternal variant was suspected based on low cycle-threshold values, and maternal genomic DNA was extracted from the original sample using the QIAasymphony SP automated DNA-extraction method and the QIAamp DNA Blood Midi kit (Qiagen) according to the manufacturer's instructions. The maternal DNA samples were then tested for *RHD* and *RHCE* using the SSP-PCR RBC-Ready Gene CDE kit and/or the weak D kit (inno-train Diagnostik) according to the manufacturer's instructions. The *RHD* result still remained inconclusive in 35 cases and these samples were followed up with sequencing.

### RHD zygosity testing by droplet digital PCR

All 35 maternal samples containing unidentified *RHD* variants were tested for *RHD* zygosity using the QX100 droplet digital PCR (ddPCR) platform and QuantaSoft v1.7 software (Bio-Rad Laboratories), as described in Sillence et al.,<sup>15,16</sup> to determine whether the women carried 1 (hemizygous Dd) or 2 copies (homozygous DD) of the *RHD* gene. Primers and probes used in ddPCR zygosity testing were adopted from Fan et al.<sup>17</sup> and Finning et al.<sup>18</sup> The samples were tested for *RHD* exons 5 and 7, against the reference gene *AGO1* on chromosome 1. The number of *RHD* copies per microliter for each exon was compared with the reference gene (*AGO1*) copy number. A sample giving a ratio of  $\sim 0.5$  was considered hemizygous and a sample giving a ratio of  $\sim 1$  was considered homozygous for the *RHD* gene.

### RHD genotyping using NGS

In all 35 maternal samples containing unidentified *RHD* variants, the *RHD* gene was amplified in 6 overlapping amplicons using long-range PCR (LR-PCR) primers described in Tounsi et al.<sup>9</sup> Primers from Hyland et al.<sup>19</sup> were also used to amplify the *RHD* gene from exon 2 to exon 7 in a 22 117-bp amplicon for 1 sample that failed to amplify amplicons 1, 2, and 3 of the 6 LR-PCR *RHD* amplicons. The sequencing library was prepared as in Tounsi et al.<sup>9</sup> Amplicons were attached to ion-sphere particles using the Ion PGM Template OT2 400 kit (Life Technologies) according to the manufacturer's instructions. Multiplexed sequencing was conducted using the Ion Xpress Barcode Adapters (Life Technologies) and a 316 chip (Life Technologies) on an Ion Torrent Personal Genome Machine (Life Technologies) as per manufacturer guidelines.

### NGS data analysis

Quality control of the sequencing reads was done using Babraham Bioinformatics FastQC software<sup>20</sup> quality check. Data analysis of

Table 1. Results of the samples (n = 35) genotyped for the *RHD* gene using NGS

No. of samples	Serological RhD typing	RHCE genotype	RHD zygosity	Nucleotide change(s)	Type	Consequence	Exon (intron)	Amino acid change(s)	Allele*	ISBT group	Linked DCE haplotype†	Most likely variant DCE haplotype‡	Allele frequency,§ %	Reference
1	Neg	Cce	Hemizygous	712delG	Del	Frame shift	5	Val238fs	<i>RHD*01N.33</i>	D <sup>-</sup>	Dce	Dce	0.006	26
1	Neg	ce	Hemizygous	330_331delGT	Del	Frame shift	2	Phe111fs	<i>RHD*01N.35</i>	D <sup>-</sup>	NR	Dce	0.006	27
1	Neg	Cce	Hemizygous	952C>T	SNV	Frame shift	7	Arg318Stop	<i>RHD*01N.61</i>	D <sup>-</sup>	Dce	Dce	0.006	26
2	Neg	Cce	Hemizygous	1227G>A	SNV	Missense	9	Synonymous	<i>RHD*01EL.01</i>	DEL	Dce	Dce	0.012	28
1	Neg	Cce	Hemizygous	147delA	Del	Frame shift	1	Val50fs	<i>RHD*01EL.04</i>	DEL	Dce	Dce	0.006	26
1	Neg	Cce	Hemizygous	IVS1+6delA	Del	Intronic SNP	(1)	NA			Dce	Dce	0.006	26
1	Neg	Cce	Hemizygous	486+1G>A	Ins	Splice site change	(3)	NA	<i>RHD*01EL.08</i>	DEL	Dce	Dce	0.006	28
1	Neg	CcEe	Hemizygous	1154G>C	SNV	Missense	9	Gly385Ala	<i>RHD*01W.2</i>	Weak D	DcE	DcE	0.012	29
1	Neg													
1	Weak pos	Ce	Hemizygous	8C>G	SNV	Missense	1	Ser3Cys	<i>RHD*01W.3</i>	Weak D	Dce	Dce	0.006	29
1	Weak pos	Cce	Hemizygous	1212C>A	SNV	Missense	9	Asp404Glu	<i>RHD*01W.72</i>	Weak D	NR	Dce	0.006	30
1	Neg¶	Cce	Hemizygous	885G>T	SNV	Missense	6	Met295Ile	<i>RHD*11</i>	Partial D	Dce, Dce	Dce	0.006	29
1	Neg	Cce	Hemizygous	845G>A	SNV	Missense	6	Gly282Asp	<i>RHD*15</i>	Partial D	DcE	Dce	0.012	29
1	Pos	ce	Hemizygous	602C>G	SNV	Missense	4	Thr201Arg	<i>RHD*09.03.01</i>	Partial D	Dce	Dce	0.006	29
1				667T>G	SNV	Missense	5	Phe233Val						
1				819G>A	SNV	Missense	6	Synonymous						
1	Neg	ce	Hemizygous	395_396dupGG	Ins	Frame shift	3	Lys133fs	<i>RHD(395-396dupGG)</i>	No designation	NR	Dce	0.006	31
4	Neg	Cce	Hemizygous	829G>A	SNV	Missense	6	Gly277Arg	New	NR	NR	Dce	0.030	MN365996
1		Ce												
3	Neg	Cce	Hemizygous	421delG	Del	Frame shift	3	Val141fs	New	NR	NR	Dce	0.018	MN365995
1				422T>A	SNV	Missense	3	Val141Glu						
1	Weak pos	Cce	Hemizygous	782C>T	SNV	Missense	5	Pro261Leu	New	NR	NR	Dce	0.012	MN366000
1		Ce												
1	Neg	Cce	Hemizygous	784delC	Del	Frame shift	5	Gln262fs	New	NR	NR	Dce	0.006	MN365997
1	Neg	cEe	Hemizygous	519C>G	SNV	Nonsense	4	Tyr173Stop	New	NR	NR	DcE	0.006	MN365999
1	Weak pos	cEe	Hemizygous	1016G>C	SNV	Missense	7	Gly339Ala	New	NR	NR	Dce	0.006	MN366002
1	Neg	cEe	Hemizygous	1154G>C	SNV	Missense	9	Gly385Ala	New	NR	DcE	Dce	0.006	MN365998
1				1163T>G	SNV	Missense	9	Leu388Arg						
1	Neg	Cce	Hemizygous	8C>G	SNV	Missense	1	Ser3Cys	New	NR	Dce	Dce	0.006	MN366001
1				49delG	Del	Frame shift	1	Ala17fs						
2#	Neg	Cce	Hemizygous	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Gray indicates D<sup>-</sup> alleles; green, DEL alleles; blue, weak D alleles; orange, partial D alleles; yellow, novel variant alleles; and red, inconclusive results.

Del, deletion; Ins, insertion; NA, not applicable; Neg, negative; NR, not reported; pos/Pos, positive; SNV, single-nucleotide variation.

\*International Society of Blood Transfusion (ISBT) designation for alleles.

†Linked haplotype reported in the RhesusBase database.

‡To assign the most possible DCE haplotype, RhCcEe genotyping results, *RHD* zygosity results, and *RHD* allele determined by NGS were all considered.

§Allele frequency was calculated based on the total number of RhD<sup>-</sup> females (16253).

¶Routine RhD phenotyping was negative, but antigen titer was 1 after elution.

¶¶*RHD\*11* occurs in 2 haplotypes with distinct phenotypes. In the Dce haplotype, the phenotype is borderline weak D/DEL, whereas in the Dce haplotype, the phenotype is weak D detected in routine tests.

#Samples failed amplicon 6 in LR-PCR; thus incomplete data.

\*\*Sample failed amplicons 1, 2, and 3 in LR-PCR.

††This sample had *RHD5/AGO1* and *RHD7/AGO1* ratios of ~1 from ddPCR and was considered homozygous for the *RHD* gene. The NGS data indicated that the sample is a compound heterozygote with 2 variant *RHD* alleles.

Table 1. (continued).

No. of samples	Serological RhD typing	RHCE genotype	RHD zygosity	Nucleotide change(s)	Type	Consequence	Exon (intron)	Amino acid change(s)	Allele*	ISBT group	Linked DCE haplotype†	Most likely variant DCE haplotype‡	Allele frequency,§ %	Reference
1**	Neg	ce	Hemizygous	336_486del 602C>G 667T>G 819G>A	Del SNV SNV SNV	Frame shift Missense Missense Missense	3 4 5 6	NA Thr201Arg Phe223Val Synonymous	NA	NA	NA	Dce	NA	NA
1	Weak pos	Cce	Homozygous††	48G>C 602C>G 667T>G 819G>A	SNV SNV SNV SNV	Missense Missense Missense Missense	1 4 5 6	Trp16Cys Thr201Arg Phe223Val Synonymous	Inconclusive	NA	NA	DCE, Dce	NA	NA

Gray indicates D<sup>-</sup> alleles; green, DEL alleles; blue, weak D alleles; orange, partial D alleles; yellow, novel variant alleles; and red, inconclusive results.

Del, deletion; Ins, insertion; NA, not applicable; Neg, negative; NR, not reported; pos/Pos, positive; SNV, single-nucleotide variation.

\*International Society of Blood Transfusion (ISBT) designation for alleles.

†Linked haplotype reported in the RhesusBase database.

‡To assign the most possible DCE haplotype, RhCcEe genotyping results, RHD zygosity results, and RHD allele determined by NGS were all considered.

§Allele frequency was calculated based on the total number of RhD<sup>-</sup> females (16 253).

||Routine RhD phenotyping was negative, but antigen titer was 1 after elution.

¶RHD\*11 occurs in 2 haplotypes with distinct phenotypes. In the DCE haplotype, the phenotype is borderline weak D/DEL, whereas in the Dce haplotype, the phenotype is weak D detected in routine tests.

#Samples failed amplicon 6 in LR-PCR, thus incomplete data.

\*\*Sample failed amplicons 1, 2, and 3 in LR-PCR.

††This sample had RHD5/AGO1 and RHD7/AGO1 ratios of ~1 from ddPCR and was considered homozygous for the RHD gene. The NGS data indicated that the sample is a compound heterozygote with 2 variant RHD alleles.

the sequencing FASTQ data were performed using CLC Genomics Workbench software (version 11.0.1; Qiagen). To map the reads to the reference, the *RHCE* gene and other genes on chromosome 1 assembly were masked to prevent the reads from scattering. The mapping was performed using the default parameters with NC\_000001.11 (*Homo sapiens* chromosome 1, GRCh38.p13 Primary Assembly) as reference.<sup>21,22</sup>

In the basic variant detection, a minimum coverage of 30, a minimum count of 5, and a minimum frequency of 80 were used for the samples considered hemizygous based on the ddPCR results. For the sample containing 2 copies of the *RHD* gene, a minimum frequency of 40 was used. The single-nucleotide polymorphisms (SNPs) detected that predicted amino acid or splice site changes were compared with known *RHD* variants on RhesusBase,<sup>23</sup> Erythrogene,<sup>24</sup> and the Blood Group Antigen Facts-Book<sup>25</sup> to determine the variant *RHD* allele.

## Sanger sequencing

The PCR primers used in the verification of the novel 829G>A mutation were 5'-AGTAGTGAGCTGGCCCATCA-3' and 5'-AGCAGAGGAGGTTAGTTGTCT-3'. The latter was used as a sequencing reverse primer as well, whereas 5'-TCTAGTTTCTTACCGGCA GGT-3' was used as a sequencing forward primer. PCR was set up using AmpliTaq DNA polymerase according to the manufacturer's instructions (Thermo Fisher Scientific), using both primers at 0.2-μM concentration. The sequencing of the PCR products was performed by the DNA Sequencing and Genomics Unit of the Institute of Biotechnology, Helsinki Institute of Life Science (HiLIFE), University of Helsinki.

## Results

### RHD zygosity

ddPCR results indicated that all samples except for 1 were hemizygous for *RHD* (Table 1), showing ratios of ~0.5 for both *RHD5/AGO1* and *RHD7/AGO1*. However, 1 sample showed both *RHD5/AGO1* and *RHD7/AGO1* ratios of 1, which indicated that it carried 2 copies of the *RHD* gene (homozygous).

### Next-generation sequencing

NGS of the whole *RHD* gene revealed 21 different variant *RHD* alleles in 31 of 35 samples (Table 1). In total, 13 previously described *RHD* variants were identified in 16 samples. These variants include 3 null alleles, 3 DEL alleles, 3 weak D alleles, 3 partial weak D alleles, and 1 allele previously reported by Karnot et al<sup>31</sup> for which there is no designation by International Society of Blood Transfusion (ISBT) as found in RhesusBase.<sup>23</sup> Mutations that encode the variant alleles and the predicted amino acid changes are listed in Table 1.

In total, 8 different novel alleles were identified in 15 samples, and these are encoded by 4 novel missense mutations, 3 frame-shift mutations caused by deletions, and 1 nonsense mutation (Table 1). Reference sequences of the novel *RHD* variant alleles were submitted to GenBank and were each assigned a unique accession number, and the allele frequencies were calculated based on the number of RhD<sup>-</sup> women (16 253) included in the Finnish national program to prevent HDFN (Table 1). The novel SNP 829G>A in exon 6 that predicts amino acid change Gly277Arg was confirmed using Sanger sequencing because



the same nucleotide change in the adjacent base, 830G>A, has been reported to cause the *RHD\*01W.12* (*RHD*\*weak D type 12) allele.

Two mutations found in exon 9 (1154G>C, 1163T>G) predicting (Gly385Ala, Leu388Arg) were detected in 1 sample. The mutation 1154G>C encodes *RHD\*01W.2* but the allele was considered a novel variant due to the additional mutation detected. Similarly, 2 mutations (8C>G, 49delG) in exon 1 that predict Ser3Cys and a frame-shift Ala17fs were detected in 1 sample. The mutation 8C>G encodes *RHD\*01W.3* but the allele was considered a novel variant due to the presence of the frame-shift SNP.

In 3 samples, LR-PCR reactions of 1 or multiple *RHD* amplicons failed presumably due to a hybrid gene or a deletion preventing the annealing of the *RHD*-specific LR-PCR primers (Table 1). In 2 of these samples, the failed amplicon was amplicon 6, which includes the area between part of intron 8 to exon 10 of the *RHD* gene. In the third sample, the LR-PCR reactions of *RHD* amplicons 1, 2, and 3 failed, which cover the region between *RHD* exons 1 and 5. The LR-PCR was repeated for this sample using primers adopted from Hyland et al<sup>19</sup> to amplify the *RHD* gene from exon 2 to exon 7. The results showed a complete deletion of exon 3 (336-486del) with a partial deletion of intron 2, in addition to mutations detected in exons 4, 5, and 6 (602C>G, 667T>G, 819G>A) that predict amino acid changes (Thr201Arg, Phe223Val, silent), respectively.

In 1 sample, which showed a homozygous *RHD* gene using ddPCR, 4 mutations (48G>C, 602C>G, 667T>G, and 819G>A) were detected in exons 1, 4, 5, and 6, respectively. These mutations predict the amino acid changes Trp16Cys, Thr201Arg, Phe223Val, and a silent mutation, respectively. These changes encode *RHD\*09.04*, meaning that the other allele is a wild-type *RHD* allele *RHD\*01*. However, the presence of a wild-type *RHD* allele that produces a normal RhD protein does not agree with the weak D reactivity in serology.

## Discussion

In Finland, genotyping is not performed unless there is an indication of an *RHD* variant. Therefore, variants that cause no or very low expression of the D antigen are not detected in the serological tests. Thus, the true variation of *RHD* among the RhD<sup>+</sup> population in Finland has not been known. This study, encompassing a cohort of over 16 000 RhD<sup>+</sup> pregnant women screened for fetal *RHD* during the first 3 years of the national program to prevent HDFN, is the first comprehensive study of *RHD* genetics in the Finnish population. NGS genotyping of the cases that remained inconclusive using SSP-PCR revealed 21 different *RHD* variants, of which 8 were novel alleles likely specific to the original Finnish population. All novel alleles were found in women who had been born in Finland and who had typical Finnish names. The SNPs 829G>A (rs1456633324), 421delG (rs755478959), and 422T>A (rs759038817) are listed in genome-aggregation database<sup>32</sup> and have been detected in European (Finnish) samples, with the following frequencies: 0.00001794, 0.000008886, and 0.000008886, respectively.

### *RHD* variation in the Finnish population

The frequency of *RHD* variants varies between populations and depends on the structure and amount of admixture in the

population. The frequency of maternal variants detected in this study was 0.5%, which is similar to the frequencies detected in Denmark<sup>33</sup> and The Netherlands<sup>34</sup> with national programs. The frequency of *RHD*<sup>+</sup> variants among RhD<sup>+</sup> blood donors is 0.47% in Switzerland<sup>35</sup> and 0.21% in Germany.<sup>26</sup> The known variants detected in this study have been previously reported in different populations and thus reflect the effect of increasing immigration and admixture in the Finnish population. However, the variety and frequency of novel variants was surprisingly high and reflects the lack of knowledge of the diversity of *RHD* variants among the RhD<sup>+</sup> population in Finland. For example, Crottet et al<sup>35</sup> discovered 4 novel *RHD* alleles among 25 370 RhD<sup>+</sup> blood donors in a large national screening in Switzerland; de Paula Vendrame et al<sup>36</sup> found 6 novel alleles in a population of 1920 Brazilian RhD<sup>+</sup> blood donors.

### Novel *RHD* variants

In this study, 8 novel mutations of *RHD* were detected that likely constitute novel null, weak D, and DEL variants, although the phenotypes could not be confirmed because fresh samples of the variant carriers were not available for D antigen-quantifying tests. The novel 421delG, 784delC, and 519C>G mutations are expected to cause null phenotypes, which would be consistent with the observed D<sup>+</sup> results in the RhD phenotyping of the samples. In addition, 3 novel missense mutations were detected. The SNPs 829G>A, 782C>T, and 1016G>C all cause amino acid changes in the transmembrane part of the RhD protein and can thus be expected to cause a weak D or DEL phenotype. All 3 samples that were predicted to have the amino acid change Pro261Leu (782 C>T) or Gly339Ala (1016 G>C) showed weak reactions in RhD phenotyping, which is consistent with the weak D phenotype caused by missense mutations in the transmembrane parts of the protein.

On the other hand, all 5 samples with the 829G>A mutation predicting amino acid change Gly277Arg showed negative reactions in the phenotyping, which could indicate a DEL phenotype because missense mutations should not cause a null phenotype. The mutation 830G>A, that is, the adjacent base, causes the amino acid change Gly277Glu and the *RHD\*01W.12* variant allele, which has been associated with extremely low antigen density.<sup>37</sup> In both the 829G>A and 830G>A mutations, a nonpolar amino acid (glycine) is changed either to a positively charged (arginine; 829 G>A) or negatively charged (glutamic acid; 830 G>A) amino acid, so the novel Gly277Arg change observed in the samples could have similar consequences to the RhD polypeptide as the Gly277Glu change in *RHD\*01W.12*. However, the possible DEL phenotype of 829G>A causing Gly277Arg would need to be confirmed with D antigen-quantifying tests.

Two novel mutations were discovered on a backbone of *RHD\*01W.2* and *RHD\*01W.3* alleles. These samples harbored not only the mutations responsible for the mentioned alleles, but also additional mutations that may explain the negative serological results of the samples. The deletion 49delG leads to a frame shift and thus probably a nonfunctional RhD protein. The missense mutation 1163T>G predicts the amino acid change Leu388Arg, which is located in the transmembrane region of the RhD protein. Whether the phenotype in the sample is DEL could only be confirmed with antigen-quantifying tests. The frequency of

weak D types 1, 2, and 3 was ~0.2% in this study period (including also RhD<sup>+</sup> samples) and the frequency of the novel mutations on the backbone of *RHD*\*01W.2 and *RHD*\*01W.3 alleles is 0.8% (2 of 242 weak D type 1, 2, and 3 samples). Relying solely on single SNP-genotyping assays designed to detect weak D mutations, a sample could be misassigned as weak D type 2 or 3 rather than D<sup>-</sup>. However, additional mutations can cause a risk of immunization. Novel alleles are often population specific<sup>38</sup> and, thus, commercial genotyping methods based on the detection of known mutations cannot unveil them. LR-PCR NGS genotyping can improve the safety of blood transfusion practices by revealing novel mutations. When serologically weak D phenotypes are detected, laboratories should complete RhD testing by determining *RHD* genotypes because weak D type 1, 2, 3, 4.0, and 4.1 individuals can be safely transfused with D<sup>+</sup> blood without being at risk for forming alloanti-D.<sup>39,40</sup> But, if the phenotype and genotype do not correspond, the sample should be sequenced or the person managed as RhD<sup>-</sup>.

### Known *RHD* variants

Among the previously reported *RHD* variants identified in this study, the Rh-phenotyping results of the samples corresponded to the reported phenotypes of the variant alleles in all but 3 samples.

Weak D type 1, 2, and 3 are the most common *RHD* variants in White populations.<sup>29</sup> They are usually detected in the serological RhD typing and identified using the SSP-PCR method. Because women carrying weak D type 1, 2, or 3 are not considered to be at risk of immunization and do not need anti-D prophylaxis, they can be excluded from the fetal *RHD* screening. The detection of *RHD*\*01W.2 and *RHD*\*01W.3 alleles in 3 samples was thus surprising. The 3 samples identified as *RHD*\*01W.2 or *RHD*\*01W.3 had given negative or very weak reactions in the serological RhD typing and for that reason were included in the screening program.

The suppressing effect of C in the *in trans* haplotype, the Ceppellini effect,<sup>41</sup> can cause the reduction of D-antigen density and lead to a DEL phenotype in weak D samples.<sup>42,43</sup> In the 3 samples identified as *RHD*\*01W.2 and *RHD*\*01W.3 in this study, no additional mutations were found, but they all have the Ce haplotype *in trans* to the weak D allele, which could explain the decrease in D-antigen expression in these samples. These results are consistent with previous observations of a very low antigen density in weak D samples that have C *in trans*.<sup>37,44,45</sup> An antigen adsorption-elution test to quantify D-antigen expression could only be performed for 1 sample of the 3, for which a fresh blood sample was available. The adsorption-elution test indicated an antigen titer of 1, which denotes a very low antigen density and a DEL phenotype.

In the samples in which a previously reported variant was identified, the probable variant haplotypes were consistent with the reported variant haplotypes in all but 1 sample. The *RHD*\*15 allele is usually associated with antigen E and observed in the DcE variant haplotype.<sup>29</sup> However, Ye et al<sup>30</sup> observed samples that were identified as *RHD*\*15 also in the ee (Cce) phenotype. In 1 of the 2 samples identified as *RHD*\*15 in this study, the most likely variant haplotype is DCe (given RhCcEe-genotyping results, *RHD* zygosity results, and the *RHD* allele determined from NGS), which supports the observation of *RHD*\*15 also being associated with e.

### Advantage of NGS

The current methods used in *RHD* genotyping, such as SSP-PCR and real-time PCR, are insufficient to identify many variants because they require the causative mutations to be known. During the first 3 years of the national fetal *RHD* screening program in Finland, almost one-half of the detected maternal variants could not be identified (35 of 79). The SSP-PCR results, however, indicated inconsistency between phenotype and genotype and thus, a variant allele.

The results of the use of NGS in blood group genotyping to overcome limitations in other genotyping platforms have been promising.<sup>9,11,12,46</sup> Exome sequencing utilizing NGS or conventional sequencing (Sanger) could have been used to identify variants; however, exome sequencing may not have been sufficient to detect intronic variations such as intronic insertion associated with *RHD* partial alleles.<sup>12</sup> Intronic mutations detected in this study align with those in Tounsi et al<sup>9</sup> for R<sub>2</sub> and R<sub>1</sub>/R<sub>0</sub>/R<sub>2</sub> haplotypes. NGS allows the use of the LR-PCR approach, enabling the amplification of the *RHD* gene (<60 kb) in only 6 overlapping amplicons about 10 kb in size. On the other hand, Sanger sequencing cannot sequence DNA fragments larger than 2 kb.<sup>47</sup> Therefore, complete sequencing of the *RHD* gene using NGS was the preferred method, which provided a complete sequence of the gene in most samples, with high coverage that allowed accurate variant calling. The advantage of NGS over Sanger sequencing is that it has the ability to sequence the genes for multiple blood group antigens simultaneously (and not just *RHD*) and hence can be applied in routine workflows in reference laboratories looking at samples with complex variants,<sup>48</sup> as in this study.

NGS of the whole *RHD* gene in this study was efficient in detecting the exonic and intronic changes (Table 1 and data not shown) causing the variants of *RHD*, except for the hemizygous hybrid *RHD*-*CE-D* genes and/or large deletions. Hybrid genes and large deletions can cause the PCR amplicons to fail if they affect the *RHD*-specific LR-PCR primer-binding sites. In 2 samples, LR-PCR amplicon 6 covering the area from intron 8 to exon 10 of the *RHD* gene failed. In these cases of failed amplicons, other methods, like single-molecule sequencing, would be needed in order to identify the exact location and extent of the changes. In the sample in which amplicons 1, 2, and 3 covering the region between exons 1 and 5 of the *RHD* gene had failed, using a different approach with primers surrounding the failed area revealed a large deletion covering part of intron 2 and all of exon 3, and an additional mutation in exon 4 in the failed region. In the future, oligonucleotide bait enrichment may allow for the sequencing of the entire *RHD* gene (or hybrid gene) directly from genomic DNA without the need for PCR amplification.<sup>49</sup>

In 1 *RHD* homozygous sample, the variant alleles could not be determined. Four heterozygous SNPs were detected that encode *RHD*\*09.04, indicating that the other *RHD* allele is a wild-type *RHD* allele. This does not agree with weak D serological results. As serological phenotyping was weak D, it indicates that the variant carrier is a compound heterozygote for 2 different variant alleles. The other variant *RHD* allele is possibly a variant *RHD* allele that harbors a deletion or a conversion that could not be detected due to the presence of an intact copy of the *RHD* gene (*RHD*\*09.04).

This was the first comprehensive study performed in the Finnish RhD<sup>-</sup> population and revealed previously unknown allelic diversity

behind the negative RhD phenotype. NGS genotyping of the entire *RHD* gene proved to be cost-efficient in uncovering the mutations causing the *RHD* variants that could not be identified using the traditional SSP-PCR technique and revealed 8 novel variants that we believe are linked to the original Finnish population. The identification of *RHD* variants causing inconclusive typing results is essential in order to prevent immunizations in pregnancies and transfusions.

## Acknowledgments

This work was supported by King Abdulaziz University (Jeddah, Saudi Arabia).

## Authorship

Contribution: S.S. and K.H. designed the study; S.M.T. and K.H. collected the data of the samples; S.M.T., W.A.T., and M.K.

performed experiments; S.M.T., W.A.T., T.E.M., and K.H. analyzed the data and wrote the manuscript; T.E.M., N.D.A., and K.H. supervised the study; and all authors reviewed, edited, and approved the manuscript.

Conflict-of-interest disclosure: N.D.A. is a consultant for YourGene Health plc. The remaining authors declare no competing financial interests.

ORCID profiles: S.M.T., 0000-0003-3696-0612; W.A.T., 0000-0002-8536-6989; S.S., 0000-0002-8908-8605; M.K., 0000-0002-3940-5385; N.D.A., 0000-0002-0151-7904; T.E.M., 0000-0002-3463-1922; K.H., 0000-0001-8904-666X.

Correspondence: Katri Haimila, Finnish Red Cross Blood Service, Kivihaantie 7, FI-00310 Helsinki, Finland; e-mail: katri.haimila@bloodservice.fi.

## References

- de Haas M, Thurik FF, van der Ploeg CPB, et al. Sensitivity of fetal RHD screening for safe guidance of targeted anti-D immunoglobulin prophylaxis: prospective cohort study of a nationwide programme in the Netherlands. *BMJ*. 2016;355(i5789):i5789.
- Urbaniak SJ. The scientific basis of antenatal prophylaxis. *Br J Obstet Gynaecol*. 1998;105(suppl 18):11-18.
- Clausen FB. Lessons learned from the implementation of non-invasive fetal RHD screening. *Expert Rev Mol Diagn*. 2018;18(5):423-431.
- Fichou Y, Audrézet MP, Guéguen P, Le Maréchal C, Férec C. Next-generation sequencing is a credible strategy for blood group genotyping. *Br J Haematol*. 2014;167(4):554-562.
- Belsito A, Magnussen K, Napoli C. Emerging strategies of blood group genotyping for patients with hemoglobinopathies. *Transfus Apheresis Sci*. 2017;56(2):206-213.
- Avent ND, Madgett TE, Halawani AJ, et al. Next-generation sequencing: academic overkill or high-resolution routine blood group genotyping? *ISBT Sci Ser*. 2015;10(suppl 1):250-256.
- Hyland CA, Roulis EV, Schoeman EM. Developments beyond blood group serology in the genomics era. *Br J Haematol*. 2019;184(6):897-911.
- Fürst D, Tsamadou C, Neuchel C, Schrezenmeier H, Mytilineos J, Weinstock C. Next-generation sequencing technologies in blood group typing. *Transfus Med Hemother*. 2020;47(1):4-13.
- Tounsi WA, Madgett TE, Avent ND. Complete RHD next-generation sequencing: establishment of reference RHD alleles [published correction appears in *Blood Adv*. 2019;3(2):120]. *Blood Adv*. 2018;2(20):2713-2723.
- Stabentheiner S, Danzer M, Niklas N, et al. Overcoming methodical limits of standard RHD genotyping by next-generation sequencing. *Vox Sang*. 2011;100(4):381-388.
- Dezan MR, Ribeiro IH, Oliveira VB, et al. RHD and RHCE genotyping by next-generation sequencing is an effective strategy to identify molecular variants within sickle cell disease patients. *Blood Cells Mol Dis*. 2017;65:8-15.
- Chou ST, Flanagan JM, Vege S, et al. Whole-exome sequencing for RH genotyping and alloimmunization risk in children with sickle cell anemia. *Blood Adv*. 2017;1(18):1414-1422.
- Schoeman EM, Roulis EV, Liew YW, et al. Targeted exome sequencing defines novel and rare variants in complex blood group serology cases for a red blood cell reference laboratory setting. *Transfusion*. 2018;58(2):284-293.
- Haimila K, Sulin K, Kuosmanen M, et al. Targeted antenatal anti-D prophylaxis program for RhD-negative pregnant women - outcome of the first two years of a national program in Finland. *Acta Obstet Gynecol Scand*. 2017;96(10):1228-1233.
- Sillence KA, Roberts LA, Hollands HJ, et al. Fetal sex and RHD genotyping with digital PCR demonstrates greater sensitivity than real-time PCR. *Clin Chem*. 2015;61(11):1399-1407.
- Sillence KA, Halawani AJ, Tounsi WA, et al. Rapid RHD zygosity determination using digital PCR. *Clin Chem*. 2017;63(8):1388-1397.
- Fan HC, Blumenfeld YJ, El-Sayed YY, Chueh J, Quake SR. Microfluidic digital PCR enables rapid prenatal diagnosis of fetal aneuploidy. *Am J Obstet Gynecol*. 2009;200(5):543.e1-543.e7.
- Finning K, Martin P, Summers J, Massey E, Poole G, Daniels G. Effect of high throughput RHD typing of fetal DNA in maternal plasma on use of anti-RhD immunoglobulin in RhD negative pregnant women: prospective feasibility study. *BMJ*. 2008;336(7648):816-818.
- Hyland CA, Millard GM, O'Brien H, et al. Non-invasive fetal RHD genotyping for RhD negative women stratified into RHD gene deletion or variant groups: comparative accuracy using two blood collection tube types. *Pathology*. 2017;49(7):757-764.
- Babraham Institute. Babraham Bioinformatics. <https://www.bioinformatics.babraham.ac.uk/>. Accessed 1 September 2018.
- Tatusova T, DiCuccio M, Badretdin A, et al. NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res*. 2016;44(14):6614-6624.



22. Sherry ST, Ward MH, Kholodov M, et al. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* 2001;29(1):308-311.
23. Wagner FF, Flegel WA. The Rhesus site. *Transfus Med Hemother.* 2014;41(5):357-363.
24. Möller M, Jöud M, Storry JR, Olsson ML. ErythroGene: a database for in-depth analysis of the extensive variation in 36 blood group systems in the 1000 Genomes Project. *Blood Adv.* 2016;1(3):240-249.
25. Reid ME, Lomas-Francis C, Olsson ML. *The Blood Group Antigen FactsBook*. 3rd ed. Boston, MA: Academic Press; 2012.
26. Flegel WA, von Zabern I, Wagner FF. Six years' experience performing RHD genotyping to confirm D- red blood cell units in Germany for preventing anti-D immunizations. *Transfusion.* 2009;49(3):465-471.
27. Le Maréchal C, Guerry C, Benech C, et al. Identification of 12 novel RHD alleles in western France by denaturing high-performance liquid chromatography analysis. *Transfusion.* 2007;47(5):858-863.
28. Wagner FF, Frohmajer A, Flegel WA. RHD positive haplotypes in D negative Europeans. *BMC Genet.* 2001;2:10.
29. Wagner FF, Gassner C, Müller TH, Schönlitzer D, Schunter F, Flegel WA. Molecular basis of weak D phenotypes. *Blood.* 1999;93(1):385-393.
30. Ye SH, Wu DZ, Wang MN, et al. A comprehensive investigation of RHD polymorphisms in the Chinese Han population in Xi'an. *Blood Transfus.* 2014;12(3):396-404.
31. Karnot A, Woestmann SJ, Flesch BK. Weak RhD phenotype caused by compound heterozygosity for DAU-2 and a new RHD\*c393-394 dupGG-mutation. *Transfus Med Hemother.* 2016;43(S1):42-43.
32. Karczewski KJ, Franciolli LC, Tiao G, et al; Genome Aggregation Database Consortium. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature.* 2020;581(7809):434-443.
33. Clausen FB, Christiansen M, Steffensen R, et al. Report of the first nationally implemented clinical routine screening for fetal RHD in D- pregnant women to ascertain the requirement for antenatal RhD prophylaxis. *Transfusion.* 2012;52(4):752-758.
34. Stegmann TC, Veldhuisen B, Bijman R, et al. Frequency and characterization of known and novel RHD variant alleles in 37 782 Dutch D-negative pregnant women. *Br J Haematol.* 2016;173(3):469-479.
35. Crottet SL, Henny C, Meyer S, et al. Implementation of a mandatory donor RHD screening in Switzerland. *Transfus Apheresis Sci.* 2014;50(2):169-174.
36. de Paula Vendrame TA, Prisco Arnoni C, Guilhem Muniz J, et al. Characterization of RHD alleles present in serologically RHD-negative donors determined by a sensitive microplate technique. *Vox Sang.* 2019;114(8):869-875.
37. Wagner FF, Frohmajer A, Ladewig B, et al. Weak D alleles express distinct phenotypes. *Blood.* 2000;95(8):2699-2708.
38. Gassner C, Doescher A, Drnovsek TD, et al. Presence of RHD in serologically D-, C/E+ individuals: a European multicenter study. *Transfusion.* 2005;45(4):527-538.
39. Sandler SG, Flegel WA, Westhoff CM, et al; College of American Pathologists Transfusion Medicine Resource Committee Work Group. It's time to phase in RHD genotyping for patients with a serologic weak D phenotype. *Transfusion.* 2015;55(3):680-689.
40. Flegel WA, Denomme GA, Queenan JT, et al. It's time to phase out "serologic weak D phenotype" and resolve D types with RHD genotyping including weak D type 4. *Transfusion.* 2020;60(4):855-859.
41. Ceppellini R, Dunn LC, Turri M. An interaction between alleles at the Rh locus in man which weakens the reactivity of the Rh0 factor (Du). *Proc Natl Acad Sci USA.* 1955;41(5):283-288.
42. Araszkiewicz P, Szymanski IO. Quantitative studies on the Rh-antigen D. Effect of the C gene. *Transfusion.* 1987;27(3):257-261.
43. Rh and RHAG blood group systems. In: Daniels G, ed. *Human Blood Groups*, 3rd ed. Oxford, United Kingdom: Wiley-Blackwell; 2013:185.
44. Polin H, Danzer M, Gaszner W, et al. Identification of RHD alleles with the potential of anti-D immunization among seemingly D- blood donors in Upper Austria. *Transfusion.* 2009;49(4):676-681.
45. Mota M, Fonseca NL, Rodrigues A, Kutner JM, Castilho L. Anti-D alloimmunization by weak D type 1 red blood cells with a very low antigen density. *Vox Sang.* 2005;88(2):130-135.
46. Fichou Y, Mariez M, Le Maréchal C, Férec C. The experience of extended blood group genotyping by next-generation sequencing (NGS): investigation of patients with sickle-cell disease. *Vox Sang.* 2016;111(4):418-424.
47. Zhang J, Chiodini R, Badr A, Zhang G. The impact of next-generation sequencing on genomics. *J Genet Genomics.* 2011;38(3):95-109.
48. Montemayor-Garcia C, Westhoff CM. The "next generation" reference laboratory? *Transfusion.* 2018;58(2):277-279.
49. Karamitros T, Magiorkinis G. A novel method for the multiplexed target enrichment of MinION next generation sequencing libraries using PCR-generated baits. *Nucleic Acids Res.* 2015;43(22):e152-e152.